OIPE

#2 —

RAW SEQUENCE LISTING   DATE: 08/21/2001
PATENT APPLICATION: **US/09/898,238**   TIME: 08:22:03

Input Set : **A:\sequence.txt**
Output Set: **N:\CRF3\08212001\I898238.raw**

ENTERED

SEQUENCE LISTING
```
 4  (1) GENERAL INFORMATION:
 6      (i) APPLICANT: Wackett, Lawrence P.
 7                    Sadowsky, Michael J.
 8                    de Souza, Mervyn L.
10     (ii) TITLE OF INVENTION: An Isolated and Purified DNA Molecule
11                             and Protein for the Degradation of Triazine Compounds
13    (iii) NUMBER OF SEQUENCES: 2
15     (iv) CORRESPONDENCE ADDRESS:
16          (A) ADDRESSEE: Mueting, Raasch, Gebhardt & Schwappach, P.A.
17          (B) STREET: P.O. Box 581415
18          (C) CITY: Minneapolis
19          (D) STATE: MN
20          (E) COUNTRY: USA
21          (F) ZIP: 55458-1415
23      (v) COMPUTER READABLE FORM:
24          (A) MEDIUM TYPE: Floppy disk
25          (B) COMPUTER: IBM PC compatible
26          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27          (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
29     (vi) CURRENT APPLICATION DATA:
30          (A) APPLICATION NUMBER: US/09/898,238
31          (B) FILING DATE: 03-Jul-2001
32          (C) CLASSIFICATION:
34   (viii) ATTORNEY/AGENT INFORMATION:
35          (A) NAME: Mueting, Ann M.
36          (B) REGISTRATION NUMBER: 33,977
37          (C) REFERENCE/DOCKET NUMBER: 110.00230101
39     (ix) TELECOMMUNICATION INFORMATION:
40          (A) TELEPHONE: 612-305-1217
41          (B) TELEFAX: 612-305-1228
44  (2) INFORMATION FOR SEQ ID NO: 1:
46      (i) SEQUENCE CHARACTERISTICS:
47          (A) LENGTH: 1858 base pairs
48          (B) TYPE: nucleic acid
49          (C) STRANDEDNESS: single
50          (D) TOPOLOGY: linear
52     (ii) MOLECULE TYPE: DNA (genomic)
54     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
56  CTCGGGTAAC TTCTTGAGCG CGGCCACAGC AGCCTTGATC ATGAAGGCGA GCATGGTGAC        60
58  CTTGACGCCG CTCTTTTCGT TCTCTTTGTT GAACTGCACG CGAAAGGCTT CCAGGTCGGT       120
60  GATGTCCGCG TCGTCGTGGT TGGTGACGTG CGGGATGACC ACCCAGTTGC GGTGCAGGTT       180
62  TTTCGATGGC ATAATATCTG CGTTGCGACG TGTAACACAC TATTGGAGAC ATATCATGCA       240
64  AACGCTCAGC ATCCAGCACG GTACCCTCGT CACGATGGAT CAGTACCGCA GAGTCCTTGG       300
66  GGATAGCTGG GTTCACGTGC AGGATGGACG GATCGTCGCG CTCGGAGTGC ACGCCGAGTC       360
68  GGTGCCTCCG CCAGCGGATC GGGTGATCGA TGCACGCGGC AAGGTCGTGT TACCCGGTTT       420
70  CATCAATGCC CACACCCATG TGAACCAGAT CCTCCTGCGC GGAGGGCCCT CGCACGGACG       480
```
C--> 30
C--> 31

RAW SEQUENCE LISTING   DATE: 08/21/2001
PATENT APPLICATION:. US/09/898,238   TIME: 08:22:03

Input Set : A:\sequence.txt
Output Set: N:\CRF3\08212001\I898238.raw

```
 72 TCAATTCTAT GACTGGCTGT TCAACGTTGT GTATCCGGGA CAAAAGGCGA TGAGACCGGA  . 540
 74 GGACGTAGCG GTGGCGGTGA GGTTGTATTG TGCGGAAGCT GTGCGCAGCG GGATTACGAC   600
 76 GATCAACGAA AACGCCGATT CGGCCATCTA CCCAGGCAAC ATCGAGGCCG CGATGGCGGT   660
 78 CTATGGTGAG GTGGGTGTGA GGGTCGTCTA CGCCCGCATG TTCTTTGATC GGATGGACGG   720
 80 GCGCATTCAA GGGTATGTGG ACGCCTTGAA GGCTCGCTCT CCCCAAGTCG AACTGTGCTC   780
 82 GATCATGGAG GAAACGGCTG TGGCCAAAGA TCGGATCACA GCCCTGTCAG ATCAGTATCA   840
 84 TGGCACGGCA GGAGGTCGTA TATCAGTTTG GCCCGCTCCT GCCACTACCA CGGCGGTGAC   900
 86 AGTTGAAGGA ATGCGATGGG CACAAGCCTT CGCCCGTGAT CGGGCGGTAA TGTGGACGCT   960
 88 TCACATGGCG GAGAGCGATC ATGATGAGCG GATTCATGGG ATGAGTCCCG CCGAGTACAT  1020
 90 GGAGTGTTAC GGACTCTTGG ATGAGCGTCT GCAGGTCGCG CATTGCGTGT ACTTTGACCG  1080
 92 GAAGGATGTT CGGCTGCTGC ACCGCCACAA TGTGAAGGTC GCGTCGCAGG TTGTGAGCAA  1140
 94 TGCCTACCTC GGCTCAGGGG TGGCCCCCGT GCCAGAGATG GTGGAGCGCG GCATGGCCGT  1200
 96 GGGCATTGGA ACAGATAACG GGAATAGTAA TGACTCCGCA AACATGATCG GAGACATGAA  1260
 98 GTTTATGGCC CATATTCACC GCGCGGTGCA TCGGGATGCG GACGTGCTGA CCCCAGAGAA  1320
100 GATTCTTGAA ATGGCGACGA TCGATGGGGC GCGTTCGTTG GGAATGGACC ACGAGATTGG  1380
102 TTCCATCGAA ACCGGCAAGC GCGCGGACCT TATCCTGCTT GACCTGCGTC ACCTCAGACG  1440
104 ACTCTCACAT CATTTGGCGG CCACGATCGT GTTTCAGGCT TACGGCAATG AGGTGGACAC  1500
106 TGTCCTGATT GACGGAAACG TTGTGATGGA GAACCGCCGC TTGAGCTTTC TTCCCCCTGA  1560
108 ACGTGAGTTG GCGTTCCTTG AGGAAGCGCA GAGCCGCGCC ACAGCTATTT TGCAGCGGGC  1620
110 GAACATGGTG GCTAACCCAG CTTGGCGCAG CCTCTAGGAA ATGACGCCGT TGCTGCATCC  1680
112 GCCGCCCCTT GAGGAAATCG CTGCCATCTT GGCGCGGCTC GGATTGGGGG GCGGACATGA  1740
114 CCTTGATGGA TACAGAATTG CCATGAATGC GGCACTTCCG TCCTTCGCTC GTGTGGAATC  1800
116 GTTGGTAGGT GAGGGTCGAC TGCGGGCGCC AGCTTCCCGA AGAGGTGAAA GGCCCGAG    1858
124 (2) INFORMATION FOR SEQ ID NO: 2:
126       (i) SEQUENCE CHARACTERISTICS:
127           (A) LENGTH: 473 amino acids
128           (B) TYPE: amino acid
129           (C) STRANDEDNESS: single
130           (D) TOPOLOGY: linear
132      (ii) MOLECULE TYPE: protein
134      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
136       Met Gln Thr Leu Ser Ile Gln His Gly Thr Leu Val Thr Met Asp Gln
137       1               5                   10                  15
139       Tyr Arg Arg Val Leu Gly Asp Ser Trp Val His Val Gln Asp Gly Arg
140                   20                  25                  30
142       Ile Val Ala Leu Gly Val His Ala Glu Ser Val Pro Pro Pro Ala Asp
143                   35                  40                  45
145       Arg Val Ile Asp Ala Arg Gly Lys Val Val Leu Pro Gly Phe Ile Asn
146               50                  55                  60
148       Ala His Thr His Val Asn Gln Ile Leu Leu Arg Gly Gly Pro Ser His
149       65                  70                  75                  80
151       Gly Arg Gln Phe Tyr Asp Trp Leu Phe Asn Val Val Tyr Pro Gly Gln
152                       85                  90                  95
154       Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu Tyr Cys
155                   100                 105                 110
157       Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn Ala Asp
158               115                 120                 125
160       Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala Val Tyr Gly
161           130                 135                 140
```

```
163    Glu Val Gly Val Arg Val Val Tyr Ala Arg Met Phe Phe Asp Arg Met
164    145             150             155             160
166    Asp Gly Arg Ile Gln Gly Tyr Val Asp Ala Leu Lys Ala Arg Ser Pro
167                    165             170             175
169    Gln Val Glu Leu Cys Ser Ile Met Glu Glu Thr Ala Val Ala Lys Asp
170                    180             185             190
172    Arg Ile Thr Ala Leu Ser Asp Gln Tyr His Gly Thr Ala Gly Gly Arg
173                    195             200             205
175    Ile Ser Val Trp Pro Ala Pro Ala Thr Thr Thr Ala Val Thr Val Glu
176            210             215             220
178    Gly Met Arg Trp Ala Gln Ala Phe Ala Arg Asp Arg Ala Val Met Trp
179    225             230             235             240
181    Thr Leu His Met Ala Glu Ser Asp His Asp Glu Arg Ile His Gly Met
182                    245             250             255
184    Ser Pro Ala Glu Tyr Met Glu Cys Tyr Gly Leu Leu Asp Glu Arg Leu
185                    260             265             270
187    Gln Val Ala His Cys Val Tyr Phe Asp Arg Lys Asp Val Arg Leu Leu
188            275             280             285
190    His Arg His Asn Val Lys Val Ala Ser Gln Val Val Ser Asn Ala Tyr
191            290             295             300
193    Leu Gly Ser Gly Val Ala Pro Val Pro Glu Met Val Glu Arg Gly Met
194    305             310             315             320
196    Ala Val Gly Ile Gly Thr Asp Asn Gly Asn Ser Asn Asp Ser Ala Asn
197                    325             330             335
199    Met Ile Gly Asp Met Lys Phe Met Ala His Ile His Arg Ala Val His
200            340             345             350
202    Arg Asp Ala Asp Val Leu Thr Pro Glu Lys Ile Leu Glu Met Ala Thr
203            355             360             365
205    Ile Asp Gly Ala Arg Ser Leu Gly Met Asp His Glu Ile Gly Ser Ile
206            370             375             380
208    Glu Thr Gly Lys Arg Ala Asp Leu Ile Leu Leu Asp Leu Arg His Leu
209    385             390             395             400
211    Arg Arg Leu Ser His His Leu Ala Ala Thr Ile Val Phe Gln Ala Tyr
212            405             410             415
214    Gly Asn Glu Val Asp Thr Val Leu Ile Asp Gly Asn Val Val Met Glu
215            420             425             430
217    Asn Arg Arg Leu Ser Phe Leu Pro Pro Glu Arg Glu Leu Ala Phe Leu
218            435             440             445
220    Glu Glu Ala Gln Ser Arg Ala Thr Ala Ile Leu Gln Arg Ala Asn Met
221            450             455             460
223    Val Ala Asn Pro Ala Trp Arg Ser Leu
224    465             470
```

VERIFICATION SUMMARY   DATE: 08/21/2001
PATENT APPLICATION: **US/09/898,238**   TIME: 08:22:04

Input Set : **A:\sequence.txt**
Output Set: **N:\CRF3\08212001\I898238.raw**

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]